IT Unit 3

Topic 2

# Ch 3, Data Analytics: Drawing Conclusions, Part 2

**Identifying patterns and relationships between data, 152-165**

1. **What is the purpose of using statistics when working with large data sets?**

Convert large quantities of raw data into small, informative, meaningful summaries. Allow patterns & relationships to be identified and allowed them to be read much easier.

1. **List the 3 most useful statistical concepts from table 3.7, p 152.** Average, Standard Deviation and Correlation.

**Basic statistics**

1. **Distinguish between the three methods of calculating an average.** Mean is the word most often referred for averages, while mode and median refer to different calculations. Median refers to the order from the smallest number to the largest number.
2. **What is the purpose of the standard deviation?** To determine how consistent a data set is how close the values are from each other, this also works out how inconstant they are from each other. A low standard deviation means that you can trust the mean, while a high one represents you shouldn’t trust it.

**Correlation and causality, p 156**

1. **Why do you need to be careful when looking at patterns in data in terms of cause and effect?** Because sometimes the data we see doesn’t make sense alongside the analysed set. There is always a correlation between both the cause and effect, sometimes although the statistics and formulas suggest something that aren’t true.

**Data visualisations, p 157**

1. **What are data visualisations used for?** It makes it easier to determine, analysis and understand complicated data sets. They are used to create graphs, charts, spatial relationships, maps, histograms and network diagrams.

**Queries and searches, p 159**

1. **How can queries, searches, filtering and sorting be used when analysing large data sets?** They can allow the user to filter certain data much easier, allowing them to find the exact data set that they are after. Alongside searches, this allows you to find the data much easier by searching for dates, keywords etc..

**Conditional formatting, p 164**

1. **Explain the process of conditional formatting:**  Conditional formatting allows you to change the appearance of data automatically based on its current value in both spreadsheets and databases.

This highlights data of special interest and can also point out crucial errors in the data.